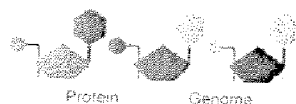


# EXHIBIT 1



My NCBI  
[Sign In]  
[Register]

Search Nucleotide

for

Go Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full)

Show 5

Send to

Range: from 3527982

to 3528476

Show whole sequence

☐ Reverse complemented str☐ 1: BA000043. Reports *Geobacillus kaust...*[gi:56378377]

Links

Features Sequence

LOCUS BA000043 495 bp DNA linear BCT 04-DEC-2004  
 DEFINITION *Geobacillus kaustophilus* HTA426 DNA, complete genome.  
 ACCESSION BA000043 REGION: 3527982..3528476  
 VERSION BA000043.1 GI:56378377  
 KEYWORDS .  
 SOURCE *Geobacillus kaustophilus* HTA426  
 ORGANISM *Geobacillus kaustophilus* HTA426  
 Bacteria; Firmicutes; Bacillales; Bacillaceae; *Geobacillus*.  
 REFERENCE 1  
 AUTHORS Takami,H., Takaki,Y., Chee,G.J., Nishi,S., Shimamura,S., Suzuki,H.,  
 Matsui,S. and Uchiyama,I.  
 TITLE Thermoadaptation trait revealed by the genome sequence of  
 thermophilic *Geobacillus kaustophilus*  
 JOURNAL (er) Nucleic Acids Res. 32 (21), 6292-6303 (2004)  
 PUBMED 15576355  
 REFERENCE 2 (bases 1 to 495)  
 AUTHORS Takami,H., Takaki,Y. and Chee,G.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-JUN-2003) Hideto Takami, Japan Marine Science and  
 Technology Center, Microbial Genome Analysis Research Group; 2-15  
 Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan  
 (E-mail:takamih@jamstec.go.jp,  
 URL:http://www.jamstec.go.jp/jamstec-e/bio/exbase.html,  
 Tel:81-46-867-9643, Fax:81-46-867-9645)  
 FEATURES  
 source Location/Qualifiers  
 1..495  
 /organism="Geobacillus kaustophilus HTA426"  
 /mol\_type="genomic DNA"  
 /strain="HTA426"  
 /isolation\_source="isolated from the deepest Ocean"  
 /db\_xref="taxon:235909"  
 /note="thermophile"  
 gene complement(1..495)  
 /locus\_tag="GK3481"  
 CDS complement(1..495)  
 /locus\_tag="GK3481"  
 /codon\_start=1  
 /transl\_table=11  
 /product="single-strand DNA-binding protein"  
 /protein\_id="BAD77766.1"  
 /db\_xref="GI:56381858"  
 /translation="MINRVILVGRLLTRDPELRYTPSGVAVATFTLAVNRPFTNQGER  
 ETDFIQCVVWRRQAENVANFLKKGSLAGVDGRLQTRSYENQEGRRVYVTEVVADSVQF"

LEPKGTSEQRGATAGGYGDPFPGQDQNHQYPNEKGFGRIDDDPFANDGQPIDISDD  
DLPF"

ORIGIN

```
1  ttaaaaacggc aaatcatcat cagaaatata gatcggtcgg cgcgcattgg cgaagagatc
61 gtoatcgatg cggccaaaacc etttttcgtt cggatattgg tggttctgat cttgcccga
121 tgggaatgga tccccatagt agcgcctcgc tgttgccctc cgcgcctcgc tgcctccttt
181 cggctcaaga aattggacgc tatcagccac cacttcgcgc acgtacacgc gccgaccttc
241 ttgattttca tagctgcggg tttgcagtcg gccatcgaca ccagccaagc tccccctttt
301 caaaaagtgg gcgacgtttt ccgcctggcg gcgccaaacg acacattgaa taaaatcgt
361 ttcccgctcg ccctgctgat ttgtaaaagg acggttgacc gcgagcgtaa acgtggcaac
421 agccactcgg cttggagtgt aacgcaactc cggatctctc gttaacctgc cgacccaaat
481 gacgcgggta atcat
```

//

Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Jul 24 2006 17:22:14



PubMed

Nucleotide



Protein

Genome

Structure



PMC

Taxonomy

OMIM

My NCBI

[Sign In]  
[Register]

Books

Search Nucleotide



for

Go Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full)

Show 5

Send to

Range: from 5398259

to 5398771

Show whole sequence

☐ Reverse complemented str☐ 1: AE016877. Reports *Bacillus cereus* A...[gi:29899096]

Links

Comment Features Sequence

LOCUS AE016877 513 bp DNA linear BCT 30-DEC-2005

DEFINITION *Bacillus cereus* ATCC 14579, complete genome.

ACCESSION AE016877 REGION: 5398259..5398771

VERSION AE016877.1 GI:29899096

KEYWORDS .

SOURCE *Bacillus cereus* ATCC 14579

ORGANISM *Bacillus cereus* ATCC 14579

Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*; *Bacillus cereus* group.

REFERENCE 1 (bases 1 to 513)

AUTHORS Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelon,B., Kapatral,V., Bhattacharyya,A., Reznik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goltsman,E., Larsen,N., D'Souza,M., Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.

TITLE Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*

JOURNAL Nature 423 (6935), 87-91 (2003)

PUBMED 12721630

REFERENCE 2 (bases 1 to 513)

AUTHORS Candelon,B., Gailloux,K., Ehrlich,D.S. and Sorokin,A.

TITLE The number of ribosomal RNA operons in *Bacillus cereus*

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 513)

AUTHORS Ivanova,N., Sorokin,A., Anderson,I., Galleron,N., Candelon,B., Kapatral,V., Bhattacharyya,A., Reznik,G., Mikhailova,N., Lapidus,A., Chu,L., Mazur,M., Goltsman,E., Larsen,N., D'Souza,M., Walunas,T., Grechkin,Y., Pusch,G., Haselkorn,R., Fonstein,M., Ehrlich,D.S.D., Overbeek,R. and Kyrpides,N.

TITLE Direct Submission

JOURNAL Submitted (12-MAR-2003) INRA, Genetique Microbienne, Domaine de Vilvert, Jouy en Josas 78352, France

COMMENT On or before Dec 30, 2005 this sequence version replaced gi:29893818, gi:29894079, gi:29894343, gi:29894619, gi:29894935, gi:29895263, gi:29895610, gi:29895852, gi:29896139, gi:29896415, gi:29896738, gi:29897026, gi:29897311, gi:29897592, gi:29897926, gi:29898221, gi:29898525, gi:29898824.

FEATURES

source Location/Qualifiers

1..513

/organism="Bacillus cereus ATCC 14579"

/mol\_type="genomic DNA"

/strain="ATCC 14579"

/db\_xref="ATCC:14579"

```

      /db_xref="taxon:226900"
      /focus
gene    complement(1..513)
        /locus_tag="BC_5475"
        /old_locus_tag="BC5475"
CDS     complement(1..513)
        /locus_tag="BC_5475"
        /old_locus_tag="BC5475"
        /codon_start=1
        /transl_table=11
        /product="Single-strand DNA binding protein"
        /protein_id="AAF12329.1"
        /db_xref="GI:29899058"
        /translation="MMNRVILVGRLTKDPLRYTPNGVAVATFTLAVNRAFANQQGER
EADFINCVIWRKQAEENVANYLKKGSLAGVDGRLQTRNYEGQDGKRVYVTEVLAESVQF
LEPRNGGGGEQRGSFNQQPSGAGFGNQGSNPFQSSNSGNQNSGFTKNDDFFSNVGQF
IDISDDDLPF"
ORIGIN
      1 ttagaacggt aaatcgctgt cggaaatgtc gatcgggtga cctacatttg aaaatggatc
      61 gtcattcttc gtaaatecag agttacottg gttacctgaa ttactagatt gaccaaattg
     121 gttagagcct tggttaccga aaccagctcc tgatgggttc tgattaaatg aaccacgttg
     181 ctccccacog ccattacgcg gctctaaaaa ttgcacgctt tccgcaagaa cttctgttac
     241 gtatacacgt ttaccatctt gtccctcgta attacgagtt tgaagacgcc catctacgcc
     301 tgctaageta ccttttttca aataatttgc tacgtttttc gcttggttac gccatattac
     361 acaattaata aagtcagctt caagctcacc ttgctgattg gcaaatgcgc gattcacagc
     421 taacgtaaaa gtagctactg cgacaccatt gggcgtgtaa cgtaagtcag ggtccttagt
     481 taaacgacca acgaggataa cagcattcat caa
//
```

Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Jul 24 2006 17:22:14



PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

Books



My NCBI

[Sign In]  
[Register]

Search Nucleotide



for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full)

Show 5

Send to

Range: from 554401

to 554931

Show whole sequence

☐ Reverse complemented str:☐ 1: NZ\_AAAC02000001. Reports *Bacillus anthracis*...[gi:65316885]

Links

Comment Features Sequence

LOCUS NZ\_AAAC02000001 531 bp DNA linear BCT 16-MAY-2005

DEFINITION *Bacillus anthracis* str. A2012 Bant\_02\_1, whole genome shotgun sequence.

ACCESSION NZ\_AAAC02000001 REGION: 554401..554931

VERSION NZ\_AAAC02000001.1 GI:65316885

KEYWORDS WGS.

SOURCE *Bacillus anthracis* str. A2012

ORGANISM *Bacillus anthracis* str. A2012  
Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*; *Bacillus cereus* group.

REFERENCE 1 (bases 1 to 531)

AUTHORS Read,T.D., Salzberg,S.L., Pop,M., Shumway,M., Umayam,L., Jiang,L., Holtzapple,E., Busch,J.D., Smith,K.L., Schupp,J.M., Solomon,D., Keim,P. and Fraser,C.M.

TITLE Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*

JOURNAL Science 296 (5575), 2028-2033 (2002)

PUBMED 12004073

REFERENCE 2 (bases 1 to 531)

AUTHORS NCBI Microbial Genomes Annotation Project.

TITLE Direct Submission

JOURNAL Submitted (16-MAY-2005) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA

COMMENT NOTE: Until 03/06/2003, this project code (NZ\_AAAC) belonged to *Burkholderia fungorum*. In order to match the code that was assigned to the GenBank version a change to the assembly-version number of the project, from 01 to 02, is necessitated.  
Protein-coding genes were predicted using GeneMark and GeneMarkHMM programs (kindly provided by M. Borodovsky). Functional annotation is based on CDD (Conserved Domain Database) and COG (Clusters of Orthologous Groups) assignments, it has not yet been subject to manual review. DNA sequence and predicted proteins are available for BLAST at [http://www.ncbi.nlm.nih.gov/sutils/genom\\_table.cgi](http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi).

FEATURES Location/Qualifiers

source 1..531  
/organism="Bacillus anthracis str. A2012"  
/mol\_type="genomic DNA"  
/strain="A2012"  
/db\_xref="taxon:191218"

gene complement(1..531)  
/locus\_tag="Bant\_01000618"

CDS complement(1..531)  
/locus\_tag="Bant\_01000618"

```
/codon_start=1
/product="COG0629: Single-stranded DNA-binding protein"
/protein_id="ZP_00390413.1"
/db_xref="GI:65317454"
/db_xref="COG:COG0629"
/translation="MVRLMNRVILVGRLLTKDPPDLRYTPNGVAVATFTLAVNRAFANQQ
GEREADFINCVIWRKQAENVANYLKKGSLAGVDCRLQTRNYEGQDGKRVYVTEVLAES
VQFLEPRNGGGEQRGSFNOQPSGAGFCNQSSNPFQSSNSGNQGNQGNQSGFTKNDDPF
SNVGQPIDISDDDLPF"
```

## ORIGIN

```
1 ttagaacggt aaatcgctgt cggaaatgtc gatcggttga cctacatttg aaaatggatc
61 gtcattcttc gtaaatccag agtttccttg gttaccttgg ttgcccgaat tactagattg
121 accaaatggg ttagagcttt ggttaccgaa accagctcct gatgggttgc gattaaatga
181 accacgttgc tccccaccgc cattacggcg ctctaaaaat tgtaacgttt cgcgaagaac
241 ttctgttaca tatacacgct taccatcttg tccctcgtaa ttaacagttt gaagacgtcc
301 atctacgect gctaagctac ctttttttaa ataatttgc acgtttttctg cttgtttacg
361 ccataattaca caattaataa agtcagcttc acgctcacct tgttgattcg caaatgcgcg
421 attcacagct aacgtaaaag tagctactgc aacaccattg ggcggtgaac gtaagtccgg
481 gtcccttagtt aaacgaccaa cgaggataac acgatttcac aatcgaacca c
```

//

Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Jul 24 2006 17:22:14



PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

Books



My NCBI

[Sign In]

[Register]

Search Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full)

Show 5

Send to

Range: from 5224259

to 5224780

Show whole sequence

☐ Reverse complemented str.☐ 1: AE017355. Reports *Bacillus thuringi...*[gi:49328240]

Links

Comment Features Sequence

LOCUS AE017355 522 bp DNA linear BCT 12-AUG-2004

DEFINITION *Bacillus thuringiensis* serovar konkukian str. 97-27, complete genome.

ACCESSION AE017355 REGION: 5224259..5224780

VERSION AE017355.1 GI:49328240

KEYWORDS .

SOURCE *Bacillus thuringiensis* serovar konkukian str. 97-27

ORGANISM *Bacillus thuringiensis* serovar konkukian str. 97-27  
Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*; *Bacillus cereus* group.

REFERENCE 1 (bases 1 to 522)

AUTHORS Brettin, T.S., Bruce, D., Challacombe, J.F., Gilna, P., Han, C., Hill, K., Hitchcock, P., Jackson, P., Keim, P., Longmire, J., Lucas, S., Okinaka, R., Richardson, P., Rubin, E. and Tice, H.

TITLE Complete genome sequence of *Bacillus thuringiensis* 97-27

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 522)

AUTHORS Brettin, T.S., Bruce, D., Challacombe, J.F., Gilna, P., Han, C., Hill, K., Hitchcock, P., Jackson, P., Keim, P., Longmire, J., Lucas, S., Okinaka, R., Richardson, P., Rubin, E. and Tice, H.

TITLE Direct Submission

JOURNAL Submitted (07-JUN-2004) Joint Genome Institute, Department of Energy, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT *Bacillus thuringiensis* 97-27 (subsp. konkukian (serotype H34)) was originally isolated from a case of severe human tissue necrosis (*Bacillus thuringiensis* subsp. konkukian (serotype H34) superinfection: Case report and experimental evidence of pathogenicity in immunosuppressed mice. Hernandez, E, Ramisse, F, Ducoureaux, J-P, Cruel, T, and Cavallo, J-D. J Clin Microbiol 1998 36(7):2138-2139). *B. thuringiensis* is indigenous to many habitats worldwide; these include soil, insects, deciduous and coniferous leaves (Prediction of insecticidal activity of *Bacillus thuringiensis* strains by polymerase chain reaction product profiles. Carozzi, NB, Kramer, VC, Warren, GW, Evola, S, and Koziel, MG. Appl Environ Microbiol. 1991 57(11):3057-61). *B. thuringiensis* is an insect pathogen that is widely used as a biopesticide in commercial agriculture. Infection of humans is unusual. The apparent pathogenic properties of *B. thuringiensis* 97-27 are very unusual for *B. thuringiensis*; unlike most *B. thuringiensis* isolates, this isolate is very closely related to *B. anthracis* based on phylogenetic analysis (Fluorescent amplified fragment length polymorphism analysis of *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* isolates. Hill, KK,



Ticknor, LO, Okinaka, RT, Asay, M, Blair, H, Bliss, KA, Laker, M, Pardington, PE, Richardson, AP, Tonks, M, Beecher, DJ, Kemp, JD, Kolsto, A-B, Wong, ACL, Keim, P, and Jackson, PJ. Appl Environ Microbiol 70(2):1068-1080. Plasmid and fosmid libraries were prepared at the Joint Genome Institute in Los Alamos (JGI-LANL), NM. Shotgun sequencing was performed at the JGI Production Genomics Facility (JGI-PGF) in Walnut Creek, CA to a coverage of 24x. Finishing was performed at JGI-LANL starting with 83 contigs and 16 scaffolds. Repetitive regions were identified, assembled and finished by manually checking paired reads close to each repeat in the assembly with consed and then making a subassembly for each repetitive region. Fifty five gaps were closed with primer walks and 16 by PCR. Gene predictions were obtained using Glimmer and tRNAs were identified using tRNAscan-SE. Basic analysis of the gene predictions was performed by comparing coding sequences against the PFam, BLOCKS and Prodom databases. Gene definitions and functional classes were added manually by a team of annotators at JGI-LANL, using BLAST results in addition to information from the basic analysis. A total of 5540 features have been annotated on the sequence record.

```
FEATURES             Location/Qualifiers
    source             1..522
                        /organism="Bacillus thuringiensis serovar konkukian str.
                        97-27"
                        /mol_type="genomic DNA"
                        /strain="97-27"
                        /serovar="konkukian"
                        /db_xref="taxon:281309"
    gene               complement(1..522)
                        /gene="ssb"
                        /locus_tag="BT9727_5154"
    CDS                complement(1..522)
                        /gene="ssb"
                        /locus_tag="BT9727_5154"
                        /inference="non-experimental evidence, no additional
                        details recorded"
                        /codon_start=1
                        /transl_table=11
                        /product="single-stranded DNA-binding protein"
                        /protein_id="AAT62672.1"
                        /db_xref="GI:49332026"
                        /translation="MMNRVILVGRLLTKDPLRYTPNGVAVATFTLAVNRAFANQQGER
                        EADFINCVIWRKQAENVANYLKKGSLAGVDGRLQTRNYEGQDGKRVYVTEVLAESVQF
                        LEPRNGGGEGQRGSEFNQQPSGAGFGNQSSNPFQSSNSGNQGNQNSGFTKNDDPFNSV
                        GQPIDISDDDLPF"
```

## ORIGIN

```
1  ttagaacggt aaatcgctgt cggaaatgtc gatcggttga cctacatttg aaaatggatc
61  gtcattcttc gtaaatccag agtttccttg gttaccttgg ttgcccgaat tactagattg
121  accaaatggg ttagagcttt ggttaccgaa accagctcct gatggttgct gattaaatga
181  accacgttgc tccccaccgc cattacgcgg ctctaaaaat tgtacgcttt ccgcaagaac
241  ttctgttaca tatcacgctt taccatcttg tccctcgtaa ttacgagttt gaagacgtcc
301  atctacgctt gctaagctac cttttttcaa ataatttgct acgtttttctg cttgtttacg
361  ccattattaca caattaataa agtcagcttc acgctcacct tgttgattcg caaatgcgcg
421  attcacagct aacgtaaaag tagctactgc aacaccattg ggcgtgtaac gtaagtccag
481  gtccttagtt aaacgaccaa cgaggataac acgattcctc aa
```

//

Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Jul 24 2006 17:22:14



My NCBI

[\[Sign In\]](#)  
[\[Register\]](#)[PubMed](#) [Nucleotide](#) [Protein](#) [Genome](#) [Structure](#) [PMC](#) [Taxonomy](#) [OMIM](#) [Books](#)

Search Nucleotide

for

[Go](#) [Clear](#)[Limits](#)[Preview/Index](#)[History](#)[Clipboard](#)[Details](#)Display [GenBank\(Full\)](#) [Show 5](#) [Send to](#)

Projection on the feature (#6)

[Show whole Sequence](#)Features: + [Refresh](#)☐ 1: NZ\_AAOY01000054. Reports *Bacillus weihenst...*[gi:89207994][Links](#)[Comment](#) [Features](#) [Sequence](#)

LOCUS NZ\_AAOY01000054 522 bp DNA linear BCT 06-MAR-2006  
DEFINITION *Bacillus weihenstephanensis* KBAB4 ctg231, whole genome shotgun sequence.  
ACCESSION NZ\_AAOY01000054 REGION: complement(5415..5936)  
VERSION NZ\_AAOY01000054.1 GI:89207994  
KEYWORDS WGS.  
SOURCE *Bacillus weihenstephanensis* KBAB4  
ORGANISM *Bacillus weihenstephanensis* KBAB4  
Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*; *Bacillus cereus* group.  
REFERENCE 1 (bases 1 to 522)  
AUTHORS Lapidus,A., Goltsman,E., Copeland,A., Lucas,S., Barry,K., Detter,J.C., Glavina del Rio,T., Hammon,N., Israni,S., Dalin,E., Tice,H., Bruce,D., Pitluck,S., Richardson,P., Auger,S., Galleron,N., Sanchis,V., Gohar,M., Broussole,V., Brillard,J., Guinebretiere,M.-H., Ehrlich,S.D., Lereclus,D., Aymerich,S., Nguen-the,C. and Sorokin,A.  
CONSRMTM US DOE Joint Genome Institute (JGI-PGF)  
TITLE Sequencing of the draft genome and assembly of *Bacillus weihenstephanensis* KBAB4  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 522)  
AUTHORS Larimer,F. and Land,M.  
CONSRMTM US DOE Joint Genome Institute (JGI-ORNL)  
TITLE Annotation of the draft genome assembly of *Bacillus weihenstephanensis* KBAB4  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 522)  
AUTHORS Lapidus,A., Goltsman,E., Copeland,A., Lucas,S., Barry,K., Detter,J.C., Glavina del Rio,T., Hammon,N., Israni,S., Dalin,E., Tice,H., Bruce,D., Pitluck,S., Richardson,P., Auger,S., Galleron,N., Sanchis,V., Gohar,M., Broussole,V., Brillard,J., Guinebretiere,M.-H., Ehrlich,S.D., Lereclus,D., Aymerich,S., Nguen-the,C. and Sorokin,A.  
CONSRMTM US DOE Joint Genome Institute (JGI-PGF)  
TITLE Direct Submission  
JOURNAL Submitted (02-MAR-2006) US DOE Joint Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA  
COMMENT URL -- <http://www.jgi.doe.gov>  
Contact: Paul Richardson ([microbes@cuba.jgi-psf.org](mailto:microbes@cuba.jgi-psf.org))  
Draft sequencing done at US DOE Joint Genome Institute  
Source DNA and bacteria available from Alexei Sorokin ([alexei.sorokine@jouy.inra.fr](mailto:alexei.sorokine@jouy.inra.fr))

The JGI and collaborators endorse the principles for the distribution and use of large scale sequencing data adopted by the larger genome sequencing community and urge users of this data to follow them. It is our intention to publish the work of this project in a timely fashion and we welcome collaborative interaction on the project and analysis.

(<http://www.genome.gov/page.cfm?pageID=10506376>)

Notes:

*Bacillus weihenstephanensis* KBAB4 was originally isolated from forest soil near Versailles, France, and was originally identified as *Bacillus cereus* (Vilas-Boas et al, Appl Env Microbiol, 2002, 68, 1414). More detailed phylogenetic analysis, using MLST, of the Versailles Collection of *B. cereus* and *B. thuringiensis* strains revealed that the strain grows at low temperature (6°C) and clusters with many other strains able to grow at low temperatures, including the independently isolated and characterized strains WSBC10204 and WSBC10206 (Sorokin et al, Appl Env Microbiol, 2006, 72, 1569). The latter two strains are the type strains of a newly recognized species in the *B. cereus* group able to grow in cold and having the species name *Bacillus weihenstephanensis* (Lechner et al, Int. J. Syst. Bacteriol, 1998, 48, 1373). Since, by MLST, KBAB4 is very closely related to WSBC10204 and represents many other psychrotrophic strains, it was assigned a species name *Bacillus weihenstephanensis* and the strain KBAB4 should also be considered as a type representative of this species.

FEATURES

source

Location/Qualifiers

1..522

/organism="Bacillus weihenstephanensis KBAB4"

/mol\_type="genomic DNA"

/strain="KBAB4"

/db\_xref="taxon:315730"

gene

1..522

/locus\_tag="BcerKBAB4DRAFT\_2050"

CDS

1..522

/locus\_tag="BcerKBAB4DRAFT\_2050"

/codon\_start=1

/transl\_table=11

/product="Single-strand binding protein"

/protein\_id="ZP\_01186525.1"

/db\_xref="GI:89208000"

/db\_xref="InterPro:IPR000424"

/db\_xref="InterPro:IPR004365"

/db\_xref="InterPro:IPR010913"

/db\_xref="InterPro:IPR011344"

/translation="MMNRVILVGRLLTKDPLRYTPNGVAVATFTLAVNRAFANQQGER  
EADFNCVIVRKQAEENVANYLKKGSLAGVDGRLQTRNYDGGQDKRVYVTEVLAESVQF  
LEPRNGGGEQGRGSEFNQQPSGAGFGNQSSNPFQSSNSGNQGNQGNQSGFTKNDPFSNV  
GQPIDISDDDLFF"

ORIGIN

```

1 ttgatgaatc gtgttatcct cgtttggtcgt ttaactaagg accctgactt acgttacacg
61 cccaatgggtg ttgcagtagc tactttttacg ttagctgtga atcggcatt tgcgaatcaa
121 caagggtgagc gtgaagctga ctttattaat tgtgtaatat ggcgtaaaca agcagaaaac
181 gtggcaaat atttgaaaaa aggtagctta gcaggcgtag acggacgtct tcaaaactcg
241 aattacgatg gacaagatgg taaaagtgtg tatgtaacag aagttcttgc ggagagcgta
301 caatttttag agccgcgtaa tggcgggtgg gagcaacgtg gttcattcaa tcagcaacca
361 tcaggagctg gtttcggtaa ccaaagctct aacccatttg gtcaatctag taattcaggt
421 aaccaaggta atcaaggtaa ctctggattt acgaagaatg acgatccatt ttogaatgtt
481 ggtaaccgga ttgacatttc ggacgacgat ttaccatttt aa

```

//

Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Jul 24 2006 17:22:14



My NCBI  
[Sign In]  
[Register]  
Books

Search Nucleotide

for

Go Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full) Show 5 Send to

Projection on the feature (#184)

Show whole Sequence

Features: + Refresh

☐ 1: Z99124. Reports *Bacillus subtilis*...[gi:32468840]

Links

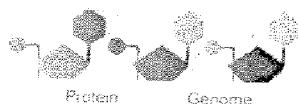
Comment Features Sequence

LOCUS BSUB0021 519 bp DNA linear BCT 18-APR-2005  
 DEFINITION *Bacillus subtilis* complete genome (section 21 of 21): from 4010730 to 4214630.  
 ACCESSION Z99124 REGION: complement(187181..187699)  
 VERSION Z99124.2 GI:32468840  
 KEYWORDS .  
 SOURCE *Bacillus subtilis* subsp. *subtilis* str. 168  
 ORGANISM *Bacillus subtilis* subsp. *subtilis* str. 168  
 Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.  
 REFERENCE 1 (bases 1 to 519)  
 AUTHORS Kunst,F., Ogasawara,N., Moszer,I., Albertini,A.M., Alloni,G., Azevedo,V., Bertero,M.G., Bessieres,P., Bolotin,A., Borchert,S., Borriass,R., Boursier,L., Brans,A., Braun,M., Brignell,S.C., Bron,S., Brouillet,S., Bruschi,C.V., Caldwell,B., Capuano,V., Carter,N.M., Choi,S.K., Codani,J.J., Connerton,I.F., Cummings,N.J., Daniel,R.A., Denizot,F., Devine,K.M., Dusterhoft,A., Ehrlich,S.D., Emmerson,P.T., Entian,K.D., Errington,J., Fabret,C., Ferrari,E., Foulger,D., Fritz,C., Fujita,M., Fujita,Y., Fuma,S., Galizzi,A., Galleron,N., Ghim,S.Y., Glaser,P., Goffeau,A., Golightly,E.J., Grandi,G., Guiseppe,G., Guy,B.J., Haga,K., Haiech,J., Harwood,C.R., Henaut,A., Hilbert,H., Holsappel,S., Hosono,S., Hullo,M.F., Itaya,M., Jones,L., Joris,B., Karamata,D., Kasahara,Y., Klaerr-Blanchard,M., Klein,C., Kobayashi,Y., Koetter,P., Koningstein,G., Krogh,S., Kumano,M., Kurita,K., Lapidus,A., Lardinois,S., Lauber,J., Lazarevic,V., Lee,S.M., Levine,A., Liu,H., Masuda,S., Mauel,C., Medigue,C., Medina,N., Mellado,R.P., Mizuno,M., Moestl,D., Nakai,S., Noback,M., Noone,D., O'Reilly,M., Ogawa,K., Ogiwara,A., Oudega,B., Park,S.H., Parro,V., Pohl,T.M., Portetelle,D., Porwollik,S., Prescott,A.M., Presecan,E., Pujic,P., Purnelle,B., Rapoport,G., Rey,M., Reynolds,S., Rieger,M., Rivolta,C., Rocha,E., Roche,B., Rose,M., Sadaie,Y., Sato,T., Scanlan,E., Schleich,S., Schroeter,R., Scoffone,F., Sekiguchi,J., Sekowska,A., Seror,S.J., Serron,P., Shin,B.S., Soldo,B., Sorokin,A., Tacconi,E., Takagi,T., Takahashi,H., Takemaru,K., Takeuchi,M., Tamakoshi,A., Tanaka,T., Terpstra,P., Tognoni,A., Tosato,V., Uchiyama,S., Vandenbol,M., Vannier,F., Vassarotti,A., Viari,A., Wambutt,R., Wedler,E., Wedler,H., Weitzenegger,T., Winters,P., Wipat,A., Yamamoto,H., Yamane,K., Yasumoto,K., Yata,K., Yoshida,K., Yoshikawa,H.F., Zumstein,E., Yoshikawa,H. and Danchin,A.  
 TITLE The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*  
 JOURNAL Nature 390 (6657), 249-256 (1997)

PUBMED 9384377  
REFERENCE 2 (bases 1 to 519)  
AUTHORS Kunst,F., Ogasawara,N., Yoshikawa,H. and Danchin,A.  
TITLE Direct Submission  
JOURNAL Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur, Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr, adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45 68 89 48  
COMMENT On Jul 7, 2003 this sequence version replaced gi:2636442. This entry contains data from release R16.1 of the Subtilist database. Further data on gene annotation and detailed information about changes from previous releases can be found at <http://genolist.pasteur.fr/Subtilist/>.  
FEATURES  
    Location/Qualifiers  
    source 1..519  
        /organism="Bacillus subtilis subsp. subtilis str. 168"  
        /mol\_type="genomic DNA"  
        /strain="168"  
    gene 1..519  
        /gene="ssb"  
        /locus\_tag="BSU40900"  
    CDS 1..519  
        /gene="ssb"  
        /locus\_tag="BSU40900"  
        /codon\_start=1  
        /transl\_table=11  
        /product="single-strand DNA-binding protein"  
        /protein\_id="CAB16127.1"  
        /db\_xref="GI:2636637"  
        /db\_xref="GOA:P37455"  
        /db\_xref="InterPro:IPR000424"  
        /db\_xref="InterPro:IPR008994"  
        /db\_xref="InterPro:IPR010913"  
        /db\_xref="Subtilist:BG10048"  
        /db\_xref="UniProtKB/Swiss-Prot:P37455"  
        /translation="MLNRVVLVGRLTKDPELRYTPNGAAVATFTLAVNRTFTNQSGER  
EADFVNCVTWRRQAENVANFLKKGSLAGVDGRLQTRNYENQQGQRVFVTEVQAESVQF  
LEPKNGGGSGSGGYNEGNSGGGQYFGGGQNDNPFPGGNQNNQRRNQNSFNDDPFANDG  
KPIDISDDDLFF"  
ORIGIN  
    1 atgcttaacc gagttgtatt agtcoggaaga ctgacaaaag acccagagct tcgttatacc  
    61 ccaaacgggtg cggtctgttc tacgtttact cttgctgtga atcgtaacatt taogaaccag  
    121 tccggagaaac gtgaggccga ttcatataat tgtgtcaact ggagaagaca agccgaaaaac  
    181 gttgcaaaact tcttgaaaaa aggaagcctt gcaggcgtag atggccggtt acaaacaaaga  
    241 aactatgaaa accagcaagg acagcgtgtc ttcgtgacag aggtccaagc tgaaagtgtt  
    301 caattttctt agccgaaaaa cggcggcgggt tctggttcag gtggatacaa cgaaggaaac  
    361 agcggcggag gccagtaact tggcggaggg caaaatgata atccatttgg gggaaatcaa  
    421 aacaaccaga gacgcaatca gggaacagc tttaatgatg acccatttgc caacgacggc  
    481 aaaccgattg acatctcgga tgatgatctt ccattctaa  
//

Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Jul 24 2006 17:22:14



My NCBI

[Sign In]  
[Register]

PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

Books

Search Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full)

Show 5

Send to

Range: from 4206965

to 4207477

Show whole sequence

☐ Reverse complemented str☐ 1: CP000002. Reports *Bacillus lichenif...*[gi:56160984]

Links

Comment Features Sequence

LOCUS CP000002 513 bp DNA linear BCT 03-DEC-2004

DEFINITION *Bacillus licheniformis* ATCC 14580, complete genome.

ACCESSION CP000002 REGION: 4206965..4207477

VERSION CP000002.2 GI:56160984

KEYWORDS .

SOURCE *Bacillus licheniformis* ATCC 14580 (DSM 13)

ORGANISM *Bacillus licheniformis* ATCC 14580

Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.

REFERENCE 1 (bases 1 to 513)

AUTHORS Rey,M.W., Ramaiya,P., Nelson,B.A., Brody-Karpin,S.D., Zaretsky,E.J., Tang,M., de Leon,A.L., Xiang,H., Gusti,V., Clausen,I.G., Olsen,P.B., Rasmussen,M.D., Andersen,J.T., Jorgensen,P.L., Larsen,T.S., Sorokin,A., Bolotin,A., Lapidus,A., Galleron,N., Ehrlich,S.D. and Berka,R.M.

TITLE Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species

JOURNAL *Genome Biol.* 5 (10), R77 (2004)

PUBMED 15461803

REFERENCE 2 (bases 1 to 513)

AUTHORS Berka,R.M., Rey,M.W. and Ramaiya,P.

TITLE Direct Submission

JOURNAL Submitted (14-JUL-2004) Novozymes Biotech Inc, 1445 Drew Ave, Davis, CA 95616, USA

REFERENCE 3 (bases 1 to 513)

AUTHORS Berka,R.M., Rey,M.W. and Ramaiya,P.

TITLE Direct Submission

JOURNAL Submitted (29-SEP-2004) Novozymes Biotech Inc, 1445 Drew Ave, Davis, CA 95616, USA

REMARK Sequence update by submitter

COMMENT On Dec 1, 2004 this sequence version replaced gi:52001702.

FEATURES

Location/Qualifiers

source 1..513

/organism="Bacillus licheniformis ATCC 14580"

/mol\_type="genomic DNA"

/strain="ATCC 14580; DSM 13"

/db\_xref="ATCC:14580"

/db\_xref="taxon:279010"

/focus

gene complement(1..513)

/gene="ssb"

/locus\_tag="BL00099"

CDS complement(1..513)

/gene="ssb"



```
/locus_tag="BL00099"  
/note="InterProMatches:IPR000424; Molecular Function:  
single-stranded DNA binding (GO:0003697)"  
/codon_start=1  
/transl_table=11  
/product="single-strand DNA-binding protein"  
/protein_id="AA025788.1"  
/db_xref="GI:52005846"  
/translation="MLNRVVLVGRLLTKDPELRYTPSGAAVATFTLAVNRTFTNQQGER  
EADFINCVVWRRQAENVANFLKKGSLAGVDGRLQTRSYENQQGQRVYVTEVQAESVQF  
LEPKGGGSGSGGYSGGQGGQHFGGGQNEPAPFGGSQNNQNRNQGNSFNDDPFANDGKP  
IDISDDDLPF"
```

## ORIGIN

```
1 ttaaaatggc aagtcgtcat cagagatgtc aatcgggttt ccatcatttg ogaatggatc  
61 gtcattgaag ctgttgccctt gattgcgatt ctgattgttt tgactgccgc ogaatggagc  
121 cggttcattt tgtccgcgcg caaagtgttg gctcccttg cctccgtgtt aacgcgcgga  
181 accagaaccg ccgccttttag gctccagaaa ttgaacactt tcagcttgaa cttccgtgac  
241 gtatacacgc tgtccctgct gattttcata gctgcgcggt tgcaaacgtc catctacacc  
301 cgcaagactt ccttttttaa ggaaatttgc aacgttttgc gcttgcttcc tccagacaac  
361 acagttgatg aaatcagctt cagcttcacc ctgctgattc gtaaacgtac gattgacagc  
421 aagcgtaaag gtggcaacag ctgcacgcgt aggagtataa cgaagctctg ggtcctttgt  
481 cagtctgccg actaaaacaa ctcggttaag cat
```

//

Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Jul 24 2006 17:22:14



PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

Books

My NCBI  
[Sign In]  
[Register]

Search Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full)

Show 5

Send to

Range: from 4186205

to 4186711

Show whole sequence

☐ Reverse complemented str:☐ 1: BA000004. Reports *Bacillus halodurans*...[gi:47118318]

Links

Comment Features Sequence

LOCUS BA000004 507 bp DNA linear BCT 01-DEC-2004  
 DEFINITION *Bacillus halodurans* C-125 DNA, complete genome.  
 ACCESSION BA000004 REGION: 4186205..4186711  
 VERSION BA000004.3 GI:47118318  
 KEYWORDS .  
 SOURCE *Bacillus halodurans* C-125  
 ORGANISM *Bacillus halodurans* C-125  
*Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.*  
 REFERENCE 1  
 AUTHORS Takami, H.  
 TITLE Genome analysis of facultatively alkaliphilic *Bacillus halodurans* C-125  
 JOURNAL (in) Horikoshi, K. and Tsujii, K. (Eds.);  
 EXTREMOPHILES IN DEEP-SEA ENVIRONMENTS: 249-284;  
 Springer-Verlag (1999)  
 REFERENCE 2  
 AUTHORS Takami, H. and Horikoshi, K.  
 TITLE Reidentification of facultatively alkaliphilic *Bacillus* sp. C-125 to *Bacillus halodurans*  
 JOURNAL Biosci. Biotechnol. Biochem. 63, 943-945 (1999)  
 REFERENCE 3  
 AUTHORS Takami, H., Nakasone, K., Hiramata, C., Takaki, Y., Masui, N., Fuji, F., Nakamura, Y. and Inoue, A.  
 TITLE An improved physical and genetic map of the genome of alkaliphilic *Bacillus* sp. C-125  
 JOURNAL Extremophiles 3 (1), 21-28 (1999)  
 PUBMED 10086841  
 REFERENCE 4  
 AUTHORS Takami, H., Nakasone, K., Ogasawara, N., Hiramata, C., Nakamura, Y., Masui, N., Fuji, F., Takaki, Y., Inoue, A. and Horikoshi, K.  
 TITLE Sequencing of three lambda clones from the genome of alkaliphilic *Bacillus* sp. strain C-125  
 JOURNAL Extremophiles 3 (1), 29-34 (1999)  
 PUBMED 10086842  
 REFERENCE 5  
 AUTHORS Takami, H., Takaki, Y., Nakasone, K., Hiramata, C., Inoue, A. and Horikoshi, K.  
 TITLE Sequence analysis of a 32-kb region including the major ribosomal protein gene clusters from alkaliphilic *Bacillus* sp. strain C-125  
 JOURNAL Biosci. Biotechnol. Biochem. 63 (2), 452-455 (1999)  
 PUBMED 10192928  
 REFERENCE 6  
 AUTHORS Takami, H., Masui, N., Nakasone, K. and Horikoshi, K.

TITLE Replication origin region of the chromosome of alkaliphilic  
 Bacillus halodurans C-125  
 JOURNAL Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999)  
 PUBMED 10427704  
 REFERENCE 7  
 AUTHORS Takami,H., Takaki,Y., Nakasone,K., Sakiyama,T., Maeno,G.,  
 Sasaki,R., Hiramata,C., Fuji,F. and Masui,N.  
 TITLE Genetic analysis of the chromosome of alkaliphilic Bacillus  
 halodurans C-125  
 JOURNAL Extremophiles 3 (3), 227-233 (1999)  
 PUBMED 10484179  
 REFERENCE 8  
 AUTHORS Takami,H. and Horikoshi,K.  
 TITLE Analysis of the genome of an alkaliphilic Bacillus strain from an  
 industrial point of view  
 JOURNAL Extremophiles 4 (2), 99-108 (2000)  
 PUBMED 10805564  
 REFERENCE 9  
 AUTHORS Nakasone,K., Masui,N., Takaki,Y., Sasaki,R., Maeno,G., Sakiyama,T.,  
 Hiramata,C., Fuji,F. and Takami,H.  
 TITLE Characterization and comparative study of the rrn operons of  
 alkaliphilic Bacillus halodurans C-125  
 JOURNAL Extremophiles 4 (4), 209-214 (2000)  
 PUBMED 10972189  
 REFERENCE 10  
 AUTHORS Takami,H., Nakasone,K., Takaki,Y., Maeno,G., Sasaki,R., Masui,N.,  
 Fuji,F., Hiramata,C., Nakamura,Y., Ogasawara,N., Kuhara,S. and  
 Horikoshi,K.  
 TITLE Complete genome sequence of the alkaliphilic bacterium Bacillus  
 halodurans and genomic sequence comparison with Bacillus subtilis  
 JOURNAL Nucleic Acids Res. 28 (21), 4317-4331 (2000)  
 PUBMED 11058132  
 REFERENCE 11 (bases 1 to 507)  
 AUTHORS Takami,H. and Takaki,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-MAR-2000) Hideto Takami, Japan Marine Science and  
 Technology Center, Deep-sea Microorganisms Research Group; 2-15  
 Natsushima, Yokosuka, Kanagawa 237-0061, Japan  
 (E-mail:takamih@jamstec.go.jp,  
 URL:http://www.jamstec.go.jp/jamstec-e/bio/DEEPSTAR/FResearch.html,  
 Tel:81-468-67-3895, Fax:81-468-66-6364)  
 COMMENT On or before Dec 8, 2004 this sequence version replaced  
 gi:10172612, gi:10172890, gi:10173176, gi:10173440, gi:10173727,  
 gi:10174030, gi:10174345, gi:10174613, gi:12641879, gi:10175192,  
 gi:10175500, gi:10175792, gi:10176109, gi:10176401.  
 FEATURES Location/Qualifiers  
 source 1..507  
 /organism="Bacillus halodurans C-125"  
 /mol\_type="genomic DNA"  
 /strain="C-125"  
 /db\_xref="taxon:272558"  
 /note="alkaliphile"  
 gene complement(1..507)  
 /gene="ssb"  
 CDS complement(1..507)  
 /gene="ssb"  
 /note="BH4049"  
 /codon\_start=1  
 /transl\_table=11  
 /product="single-strand DNA-binding protein (phage-related

```
protein)"
/protein_id="BAB07768.1"
/db_xref="GI:10176674"
/translation="MLNRVVLVGRLTRDFELRYTPNGVAVANFTLAVNRPFNSQQGER
EADFINCVVWRKQAENVANYLKKGSLAGVDGRIQTRSYDNNEGRRVFVTEVMAESVQF
LEPRGSQSQQGSNVDFGGGSPNNPMGGNDFGQQSQGSGRQSGGFSEDPFANDGKPID
ISDDDLFF"
```

## ORIGIN

```
1 ttagaatggo aaatcatcat oggaaatata gattggcttg ccatcatttg caaacgggto
61 ttctgagaaa ccacoggatt gacgtccgga gccaccagat tgttgccgga aatcattacc
121 gccattggg ttgttagggg agccgccaac aaagtgtgca acgtttgata ccccttggct
181 ttgtgagcgg cgcggttcaa ggaattgaac gctttctgcc atgaaccttg tcacaaagac
241 tctgoggcct tcattgttgt cataactgcg cgtttgaata cgaccatcaa caccagctag
301 acttcccttt ttcaaatagt togcacggtt ctccgcttgc ttccgcctaa cgacacagtt
361 aatgaaatct gcttcgcgct ctcttgttgg gttcgagAAC gggcggttta cagcaagtgt
421 aaagtgtgct acagcaacac cgtttggcgt gtagcgcaac tctggatcac gtgtcaaccg
481 accgactagg acgacacgat ttaacat
```

//

Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Jul 24 2006 17:22:14



PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

My NCBI

[Sign In]  
[Register]

Books

Search Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display GenBank(Full)

Show 5

Send to

Range: from 4288801

to 4289286

Show whole sequence

☐ Reverse complemented str:☐ 1: AP006627. Reports *Bacillus clausii* ...[gi:56908016]

Links

Features Sequence

LOCUS AP006627 486 bp DNA linear BCT 28-JUL-2006

DEFINITION *Bacillus clausii* KSM-K16 DNA, complete genome.

ACCESSION AP006627 REGION: 4288801..4289286

VERSION AP006627.1 GI:56908016

KEYWORDS .

SOURCE *Bacillus clausii* KSM-K16

ORGANISM *Bacillus clausii* KSM-K16

Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.

REFERENCE 1

AUTHORS Hakamada,Y., Kobayashi,T., Hitomi,J., Kawai,S. and Ito,S.

TITLE Molecular cloning and nucleotide sequence of the gene for an alkaline protease from the alkalophilic *Bacillus* sp. KSM-K16

JOURNAL J. Ferment. Bioeng. 78, 105-108 (1994)

REFERENCE 2

AUTHORS Kobayashi,T., Hakamada,Y., Adachi,S., Hitomi,J., Yoshimatsu,T., Koike,K., Kawai,S. and Ito,S.

TITLE Purification and properties of an alkaline protease from alkalophilic *Bacillus* sp. KSM-K16

JOURNAL Appl. Microbiol. Biotechnol. 43 (3), 473-481 (1995)

PUBMED 7632397

REFERENCE 3

AUTHORS Shirai,T., Suzuki,A., Yamane,T., Ashida,T., Kobayashi,T., Hitomi,J. and Ito,S.

TITLE High-resolution crystal structure of M-protease: phylogeny aided analysis of the high-alkaline adaptation mechanism

JOURNAL Protein Eng. 10 (6), 627-634 (1997)

PUBMED 9278275

REFERENCE 4

AUTHORS Sakaki,Y., Kageyama,Y., Shimamura,S., Suzuki,H., Nishi,S., Hatada,Y., Kawai,S., Ito,S. and Horikoshi,K.

TITLE The complete genome sequence of the alkaliphilic *Bacillus clausii* KSM-K16

JOURNAL Unpublished

REFERENCE 5 (bases 1 to 486)

AUTHORS Sakaki,Y., Kageyama,Y., Shimamura,S., Suzuki,H., Nishi,S., Hatada,Y., Kawai,S., Ito,S. and Horikoshi,K.

TITLE Direct Submission

JOURNAL Submitted (19-OCT-2003) Yasushi Kageyama, Kao Corporation, Biological Science Laboratories; 2606 Akabane, Ichikai-machi, Haga-gun, Tochigi 321-3497, Japan  
(E-mail:kageyama.yasushi@kao.co.jp, URL:http://www.kao.co.jp/e/, Tel:81-285-68-7516, Fax:81-285-68-7547)

FEATURES Location/Qualifiers

```
source      1..486
             /organism="Bacillus clausii KSM-K16"
             /mol_type="genomic DNA"
             /strain="KSM-K16"
             /db_xref="taxon:66692"
             /note="alkaliphile"
gene        complement(1..486)
             /gene="ssb"
             /locus_tag="ABC4106"
CDS         complement(1..486)
             /gene="ssb"
             /locus_tag="ABC4106"
             /codon_start=1
             /transl_table=11
             /product="single-strand DNA-binding protein SSB"
             /protein_id="BAD66637.1"
             /db_xref="GI:56912109"
             /translation="MLNRVVLVGLTRDPELRFTPNGVAVANFTLAVNRPFSNQOGER
EADFINCVVWRKPAENVANFLKKGSLAGVDGRVQTRSYDNNEGRRVFVTEIVAESVQF
LEPRNSQNQSSGNNPGFDQYGCAGNQSGNQGTGGSRSNGYDNDPFSNDGSIIDISDDDLF
F"
ORIGIN
    1 ttaaaatgga agatcatcgt cggaaatata aatcgaaccg tcattcgaga acgggtcggt
   61 atcatatcca ttagaacggc tgccaaccagt ttggttgcca gaaccttggt tgccagcgcc
  121 gtactgatca aaaccaggat tattcccaga ggattgggtt tgactgttgc gaggttcgag
  181 aaattggacg ctttcagcaa cgatttcctt tacaaaaacg cgcctgcctt cattattgtc
  241 atagcttcgc gtttgcacac ggcggtcgac tctgcaagg ctcccttttt taaggaaatt
  301 ggcgacgttc tcagccggtt tgcgccagac aacacagtta atgaagtcag cttctcgttc
  361 tccttgctgg ttgaaaaacg gacggttgac tgcaagcgta aaattggcta cagctacacc
  421 gtttggcgtg aatcgcaatt cagggtcacc cgttaagcgt ccgacaagaa caacacggtt
  481 taacaa
//
```

Disclaimer | Write to the Help Desk  
NCBI | NLM | NIH

Jul 24 2006 12:22:14

## EXHIBIT 2

```
#####
# Program: needle
# Rundate: Mon Jul 31 06:58:16 2006
# Align_format: srspair
# Report_file: /ebi/extserv/old-work/needle-20060731-06581567407497.output
#####

#=====
#
# Aligned_sequences: 2
# 1: SEQ175
# 2: G_kausotphilus
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 553
# Identity: 252/553 (45.6%)
# Similarity: 252/553 (45.6%)
# Gaps: 119/553 (21.5%)
# Score: 1104.5
#
#
#=====

SEQ175      1      atgattaaccgcgtc---attttggtcggcagg-----ttaa      34
      ..|...|...|...|  |||  ||...|...|...|...|  ||..
G_kausotphilu 1  ttaaaacggcaaatcatcatcagaaatatcgatcggtcgccgctcattgg      50

SEQ175      35  cgaga-gatc-----cggagttgcgttacac---tccaagcggga-----g      70
      |||..|  |||  ||...|...|...|...|  |...|...|  |
G_kausotphilu 51  cgaaaggatcgtcatcgatcgcgccaaacccttttcgttcggatattgg      100

SEQ175      71  tgg--ctg---ttgccacgtttacg-----ctcgcggtcaaccgctc-      106
      |||  |||  |||  ||...|...|  |...|...|  |...|...|
G_kausotphilu 101 tggttctgatcttgcc-cgaatgggaatggatccccatagt-agccgcct      148

SEQ175      107 -cgtttacaaatcagcaggcgagcggg--aaacggattttattcaatgt      153
      |...|...|...|  ||...|...|...|  ...|...|  |...|...|
G_kausotphilu 149 gctgttgccctc-gctgctcgctcgttcctttcggc--tcaagaaattg      195

SEQ175      154 gtcgtttggcgccgcca-----ggcg-gaaaacgtcgccaactttttgaa      197
      |...|...|...|...|  |...|...|...|...|  |||...|...|...|
G_kausotphilu 196 gacgctatcagccaccacttccgtcacgtacag-cgcccaccttcttga      244

SEQ175      198 aaaggggagcttggttg--tgtcgatggccgactgcaaaccgc--agc      243
      .....|...|  ||...|...|...|...|  ||...|  |||
G_kausotphilu 245 ttttcataget--gcgggtttgcagtcggccatcgac--accagccaagc      290

SEQ175      244 tatgaaaatcaagaaggtcggtg-gtgtacgtgacggaagtgggtgctg      292
      |.....|...|...|...|...|  ||...|  |...|  |||...|...|
G_kausotphilu 291 tcccccttttcaaaaagttggcgacgttttc-cgcc-----tggcggcgc      334

SEQ175      293 atagcgtccaattttcttgag--ccgaaaggaacgagcgagcagcagggg      340
      ..|...|...|...|...|  ||...|  ...|...|...|...|  |...|
G_kausotphilu 335 caaacgacacattgaataaaatcgttt--cccgctcgccctgc-tgatt      381

SEQ175      341 cgacagcagcgcggt-actatggggatccattcccattcgggcaagatca      389
      |...|...|...|...|  ||...|  |...|...|  |||...|  ||
G_kausotphilu 382 tgtaaacggacggttgaccgcgag---cgtaaacgt---ggcaa---ca      421
```



```

SEQ175      390 gaaccaccaatatccgaacgaaaaagggtttgccgcgcgcgatgaacgac 439
      |  ||||  ||||...|  ||||  |||||  ||||...
G_kausotphilu 422 g--ccac-----tccgcttg----gagtgt-aacgcaac---tccggat 455
SEQ175      440 ctttcgccaatgaaggccagccgatcgatattttgatgatgatgttgcgc 489
      ||.||||.||  ||.|||||.||.|||||  ||.||||  |.
G_kausotphilu 456 ctctcggttaa-----cctgcgcgacccaaaatgac-gcgggttaat----ca 494
SEQ175      490 ttt      492
      |
G_kausotphilu 495 t      495

```

```

#####
# Program: needle
# Rundate: Mon Jul 31 07:12:56 2006
# Align_format: srspair
# Report_file: /ebi/extserv/old-work/needle-20060731-07125510979774.output
#####

#=====
#
# Aligned_sequences: 2
# 1: SEQ175
# 2: B_cereus
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 553
# Identity: 246/553 (44.5%)
# Similarity: 246/553 (44.5%)
# Gaps: 101/553 (18.3%)
# Score: 1047.0
#
#=====

SEQ175      1      atgattaaccg-cgtcattttggtcggcaggttaacgagagatcc      44
              |.|.|.|.|.| ||||.....| |||...| |||..| |..|.|..
B_cereus    1      ttagaacggtaaatcgctcgtcggaatgtcgatcggttgac-ctacattt      49

SEQ175      45      ggagttg---cgttacaactcc---aagcggagtggt---gttgccacg      84
              |.|..| | |.|..|.| | |..| | |..| | |..| | |..
B_cereus    50      gaaaatggatcgctcattcttcgtaaaatccagagttaccttggttacctga      99

SEQ175      85      tttaacgctcgcggtaaccg-----tcc---gtttacaaatcagcagg      124
              .| | | |.....|..| | | | | | | | | | | | | | | | | |
B_cereus    100     attactagattgaccaaattgggttagagccttggttacggaaac--cagc      147

SEQ175      125     gcgagcgggaaacggattttattcaatgtgtcgtttggcgccgcca----      170
              .|..|..|..|..| | | | | | | | | | | | | | | | | | |
B_cereus    148     tcctgatggttgctgattaaatgaaccacgttgctccccaccgcccattac      197

SEQ175      171     -ggcggaaaaacgtcgccaactttttgaaaaagggagccttggtg----      214
              | | | | | | | | | | | | | | | | | | | | | | | |
B_cereus    198     gcggctctaaaaattgcacgctttccgcaa----gaactt--ctgttac      240

SEQ175      215     gtgt---cgatggccgactgcaaaccgcagctatgaaaatcaagaaggt      261
              | | | | | | | | | | | | | | | | | | | | | | | |
B_cereus    241     gtataacggtttaccatcttgctcctcgtaattacgagtttgaagacgcc      290

SEQ175      262     cggcgtgtgtacgtgacggaagtgggtggtgatagcgtccaattttcttga      311
              |..| | | | | | | | | | | | | | | | | | | | | | |
B_cereus    291     catc----tac--gcctgctaagctaccttttt--tcaataaatttgc      330

SEQ175      312     gccgaaaggaaacgagcgcagcagcagggggcgaca-----gcaggcggct      355
              ..| | | | | | | | | | | | | | | | | | | | | | |
B_cereus    331     tacgttttctgcttgttttaogccatattac-acaattaataaagtcagct      379

SEQ175      356     actatggggatccattccattcgggcaagatc-agaaccaccaatatcc      404
              |..| | | | | | | | | | | | | | | | | | | | | |
B_cereus    380     tcac--gctcaccttgctgatt--ggcaaatgcgcgattcacag-----c      420

```

```

SEQ175      405 gaacgaaaaagggtttggccgcgcgatgacgatccctttgccaatgacg      454
      .||||.|||||      |.|||.|||      ||| |.|||.|||.|||.|||
B_cereus    421 taacgtaaaag---tagctactgc---gac-accattggcggtgtaacg      462
      .||||.|||||.|||      |.|||.|||.|||.|||.|||.|||.|||
SEQ175      455 ---gccagccgatcgatat---ctgatgatgatttgccggtt      492
      |.||| |.|||.|||.|||      |.|||.|||.|||.|||.|||.|||
B_cereus    463 taagtcag--ggtccttagttaaacgaccaacgaggataacacgattcat      510
      |.||| |.|||.|||.|||      |.|||.|||.|||.|||.|||.|||
SEQ175      493      492
B_cereus    511 caa      513

```



```

SEQ175      384 ag---atcagaaccaccaatatccg---aacgaaa--aagggt----- 418
      || ..|.|.|.|||||.|||.||| |||||.|||.|||||
B_anthraxis 439 agtagctactgcaacaccattggggcgtgtaacgtaagtcagggtccttag 488

SEQ175      419 ttggccg--catcgatgacgatcc-tttcgccaatgacggccagccgatac 465
      |||.||| |||||.|||.|||.|||.|||||.|||.|||||.|||.|||.
B_anthraxis 489 ttaaacgaccaacgaggataaacacgattcatcaat--cgaaccac 531

SEQ175      466 gatatttctgatgatgatttgccgttt 492
B_anthraxis 532
      531

```

```
#####
# Program: needle
# Rundate: Mon Jul 31 07:15:58 2006
# Align_format: srepair
# Report_file: /ebi/extraerv/old-work/needle-20060731-07155589573420.output
#####
```

```
#=====
#
# Aligned_sequences: 2
# 1: SEQ175
# 2: B_thuringiensis
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 556
# Identity: 238/556 (42.8%)
# Similarity: 238/556 (42.8%)
# Gaps: 98/556 (17.6%)
# Score: 1029.5
#
#=====
```

SEQ175	1	atgattaacgc-cgtcatttttggtcggcagggttaacgagagatcc	44
B_thuringiens	1	ttagaacggtaaatcgtcgtcggaatgtcgatcggttgac-ctacattt	49
SEQ175	45	ggagttg---cgttacactcc---aagcggagtggt---gttgccacg	84
B_thuringiens	50	gaaaatggatcgtcattcttcgtaaaatccagagtttcttggttaccttg	99
SEQ175	85	tttacgc-----tcggcgtaacc-----gtcc---gtttacaa	115
B_thuringiens	100	gttgcccgaaattactagattgaccaaattgggttagagctttggttacga	149
SEQ175	116	atcagcagggcgagcgggaaacggattttattcaatgtgtcgtttggcgc	165
B_thuringiens	150	aac--cagctcctgatggttgctgattaaatgaaccagttgctcccccac	197
SEQ175	166	cgcca-----ggcggaacgcgcgaactttttgaaaaagggagctt	209
B_thuringiens	198	cgccattacgcggctctaaaaattgtacgctttccgcaa----gaactt	242
SEQ175	210	ggctggtgtcgatggccgactgcaaaccgcagctatgaaaatcaagaag	259
B_thuringiens	243	--ctgttacatatacacgcttaccatcttgcctcgttaattacgagttt	290
SEQ175	260	gtcggcgtgtgtacgtgacggaagtgtggtgatagcgtccaaatttctt	309
B_thuringiens	291	gaagacgtccatctacgcctgctaagctaccttttt---tcaaataattt	337
SEQ175	310	gagccgaaaggaacgagcagcagcagggcgaca-----gcaggcgg	353
B_thuringiens	338	gtacggtttctgctgtttacgccatattac-acaattaataaagtcag	386
SEQ175	354	ctactatggggatccattccattcgggcaagatcagaaccaaccaatc	403
B_thuringiens	387	cttcac--gtcaccttggtgattcg-----caaatgcgcgattcac	426



```

#####
# Program: needle
# Rundate: Mon Jul 31 07:18:04 2006
# Align_format: srspair
# Report_file: /ebi/extserv/old-work/needle-20060731-07180336868657.output
#####

#=====
#
# Aligned_sequences: 2
# 1: SEQ175
# 2: B_weihenstephanensis
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 525
# Identity: 334/525 (63.6%)
# Similarity: 334/525 (63.6%)
# Gaps: 36/525 ( 6.9%)
# Score: 1725.0
#
#=====

SEQ175      1 atgattaaccgcgtcatttttggtcggcagggttaacgagagatccggaggtt    50
               .|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
B_weihensteph 1 ttgatgaatcgtgttatcctcgttggtcgtttaactaaggaccctgactt    50

SEQ175      51 gcgttacactccaagcggagtggtgttgccacgtttacgctcgcggtca    100
               .|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
B_weihensteph 51 acgttacacgccaatgggtgtgcagtagctacttttacgttagctgtga    100

SEQ175      101 accgtccggtttacaaatcagcagggcgagcgggaaacggattttattcaa    150
               |.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
B_weihensteph 101 atcgcgcatttgcgaaatcaacaaggtagcgtgaagctgactttattaat    150

SEQ175      151 tgtgtcgtttggcgccgcagcgggaaaaacgtcgccaactttttgaaaaa    200
               |||.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
B_weihensteph 151 tgtgtaatatggcgtaaacagcagaaaacgtggcaaattatttgaaaaa    200

SEQ175      201 ggggagcttggtggtgtcgatggccgactgcaaaccgcagctatgaaa    250
               .|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
B_weihensteph 201 aggtagcttagcagggcgtagacggacgtcttcaaactcgtaattacgatg    250

SEQ175      251 atcaagaaggctggcgtgtgtacgtgacggaagtggtggctgatagcgtc    300
               .|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
B_weihensteph 251 gacaagatggtaaacgtgtatatgtaacagaagttcttgaggagagcgta    300

SEQ175      301 caattttcttgagccg---aaaggaacgagcgagcagcgagg-----    338
               |||.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
B_weihensteph 301 caatttttagagdcgcgtaatggcggtggggagcaacgtggttcattcaa    350

SEQ175      339 --ggcga-cagcaggcggctactatggggatcca-----ttcccatto    378
               .|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
B_weihensteph 351 tcagcaaccatcagg-agctggtttcggttaaccaagctctaaccattt    399

SEQ175      379 gggcaagatc-agaaccaccaatatccgaacgaaaaaggggtttggcgcga    427
               ||.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
B_weihensteph 400 ggtca-atctagtaattcaggtaaccaaggtaatcaaggtaactctgga    447

```







```

SEQ175          418 tttggcgcgatcgatgacgatcctttcgccaatgacggccagccgatcga      467
                  ..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
B_subtilis      445 --aacagctttaatgatgacccatttgccaacgacggccaaacgattga      491

SEQ175          468 tattctcgatgatgatttgcgcttt      492
                  .|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|
B_subtilis      492 catctcggatgatgatcttccattctaa      519

```

```
#####
# Program: needle
# Rundate: Mon Jul 31 07:20:15 2005
# Align_format: srspair
# Report_file: /ebi/extserv/old-work/needle-20060731-07201389109922.output
#####

#=====
#
# Aligned_sequences: 2
# 1: SEQ175
# 2: B_licheniformis
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 566
# Identity: 253/566 (44.7%)
# Similarity: 253/566 (44.7%)
# Gaps: 127/566 (22.4%)
# Score: 1124.0
#
#=====

SEQ175          1      atgattaaccg-cgtc---attttggtcggcaggttaacgagag      40
                |||...|...| |...| |...|...|...| |...|
B_licheniform   1      ttaaaatggcaagtgcgtcatcagagatgtcaatcggc---tttcc----      42

SEQ175          41      atccggagttgcgttacactccaagcggagtggtgttgccacgtttacg      90
                ||| |...|...|...| |...|...|...|...|...|...|
B_licheniform   43      atc---atttgcgaatggatc--gtcattgaagctgttgcccttgattgcg      87

SEQ175          91      ctgcgggtcaaccgtccgtttacaatcagcagggcgagcgggaaacgga      140
                .| |...|...|...|...|...| |...|...|...|
B_licheniform   88      at-----tctgattgtttgactgccgc---cgaat-ggagccgg-      123

SEQ175          141     ttttattcaatgtgtcgtttggcgccgcca---ggcggaaaacgtcgcc      186
                |||...|...| |...|...| |...|...| |...|...|
B_licheniform   124     -----ttcattttgtc-----cgccgccaaagtgttggcctccctggcc      162

SEQ175          187     aactttttgaaaaagggagcttggctgggtgtcgatggcgcactgca---      233
                ..|...|...|...|...|...|...|...|...|...|...|
B_licheniform   163     tccgctgtaaccgcgggaaccagaaccgcgcctttaggctccagaaatt      212

SEQ175          234     -aacc--gcagctatgaaaatcaagaaggtcggcgtg-tgtacgtgacg      279
                |||. |...|...|...|...|...|...|...|...|...|
B_licheniform   213     gaacactttcagct-tgaacttcogtgacgtatacacgctgtccctgctg      261

SEQ175          280     gaagtgggtggc-----tgatagcgtccaatttcttgagccgaaagga      321
                ....|...| |...|...|...|...|...|...|...|
B_licheniform   262     attttcatagctgcgcgtttgcaaacgtcca---tctacaccgcaag--      306

SEQ175          322     acgagc-----ga-----gcagcg---agg---ggcgacagcagg      350
                ||...| |...|...|...|...|...|...|...|...|
B_licheniform   307     acttccttttttaaggaaatttgcaacgttttcggcttgtcttctccaga      356

SEQ175          351     cggctactatgg--ggatcca--tcccattcg---ggcaagatcagaac      393
                |..| |...|...|...|...|...|...|...|...|...|
B_licheniform   357     caac-acagttgatgaaatcagcttcacgttcacccctgctgattcgtaaa      405
```

10066505.1

```
#####
# Program: needle
# Rundate: Mon Jul 31 07:21:22 2006
# Align_format: srspair
# Report_file: /ebi/extserv/old-work/needle-20060731-07212194055950.output
#####

#=====
#
# Aligned_sequences: 2
# 1: SEQ175
# 2: B_halodurans
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 553
# Identity: 253/553 (45.8%)
# Similarity: 253/553 (45.8%)
# Gaps: 107/553 (19.3%)
# Score: 1074.0
#
#=====

SEQ175      1      atgattaaccg-cgtc----attttggtcggcaggttaacg--ag      38
              |||...|||. |||  |||.|||.|||.|||.|||. |
B_halodurans 1      ttagaatggcaaatcatcatcggaatatcgattggcttgccatcatttg      50

SEQ175      39      agatccggagttgcgttacactccaagcggagtg--ctgttgccac--      83
              ..|.|||.|||.|||.|||.|||.  |||.|||.  |..| ||||
B_halodurans 51      caaacgggtcttctgagaaaccac---cggattgacgtccggagccacca      97

SEQ175      84      -gtttacgctcgcgggtca--accgtccgtttacaaatcagcagggcgagc      130
              .|||.|||.|||.|||.  |||.|||.  ...|...| ||||
B_halodurans 98      gattgttgcccgaatcattaccgcocatt---gggttggttaggg-gagc      143

SEQ175      131     gggaaacggattttattcaatgtgtcgtttggcgccgagggcggaaaac      180
              .|.|||.|||.|||.  |||.  |||.|||.|||.|||.  .....
B_halodurans 144     cgccaccaaagtt--gtcaa----cgtttgatcccccttggtttgtga      186

SEQ175      181     gtcgccaactttttgaaaaaggggagcttggtggtgtcgatggcgact      230
              |.|||.|||.|||.|||.  |||.|||.  |||.  |.|||.|||. ||
B_halodurans 187     gccgcgcggttcaaggaatt-gaacgctt-tctg---ccatgacc-tct      229

SEQ175      231     g---caaaccgcagctatgaaaatcaagaaggctg-----gcgtgtg      270
              |  |||.|||.|||.  |....|||.  |||.  |||.|||.
B_halodurans 230     gtcacaaagactctgc---ggccttcattgtgtgcataactgcgcgtttg      276

SEQ175      271     tacgtgacggaagtgggtggtgatag-cgtccaatttcttgagccgaaaag      319
              .|||.|||.  |.....|.|||.  |||.|||.  .....|.|||.
B_halodurans 277     aatccgac--catcaacaccagctagacttccctttt-----tcaaata      318

SEQ175      320     gaacg-agcgagcagcgagggggcgacag-cagggcggtac---tatgggg      364
              [..] |..|..|..|..  ..|..|.  |||.|||.  |||.  |..|..
B_halodurans 319     gttcgccacgttctccgc-ttgcttccgccaaacgac-acagttaatgaa      366

SEQ175      365     atc-----cattcccatcgggcaagatc-agaaccaccaatatccga      406
              |||  |..|.|||.  |..|.|||.  |||.|||.  |..|..
B_halodurans 367     atctgcttcgcgctctcctt---gttggttcgagaacggcggtttacag      413
```

```

SEQ175      407 acgaaaaagggtttg----gccgcgatcg-atgacgatccttctgc--ca    448
             .....|..|||||   ||..|||.||..||| |..||| |.
B_halodurans 414 caagtgtaaagtttgctacagcaaacaccggtttggcggtg--tagcgcaact    461
                               |||||
SEQ175      449 atg---acg-gccagcgcgatcgatattttct--gatg--atgatttgccg     489
             .||   |||.|||.|||   ||  ||.| |.|||||.|.
B_halodurans 462 ctggatcacgtgtccaacgcaccga-----ctaggacgacacgatttaaca    506
                                     |||||
SEQ175      490 ttt          492
             |
B_halodurans 507 t          507

```

```

#####
# Program: needle
# Rundate: Mon Jul 31 07:22:20 2006
# Align_format: srspair
# Report_file: /ebi/extserv/old-work/needle-20060731-07222012434590.output
#####

#=====
#
# Aligned_sequences: 2
# 1: SEQ175
# 2: B_clausii
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 552
# Identity: 257/552 (46.6%)
# Similarity: 257/552 (46.6%)
# Gaps: 126/552 (22.8%)
# Score: 1055.5
#
#=====

SEQ175      1      at-gattaaccgcgtc---attttggtcg-gcaggttaacgaga      39
              || ||..||..||| ||..||| ..|.....|
B_clausii   1 Ttaaaatggaagatcatcgctcggaatatcaatcgaaccgctattcgaga      50

SEQ175      40 gatccggagtttgcgtta-cactccaagcggagtggctgttgccacgttta      88
              .||| || ||||| ||...|...|..|||...||| |||
B_clausii   51 ---acgg-gt--cgttatcatatccattagaacggctgccacca-gttt-      92

SEQ175      89 cgctcgcggtc--aacgcgtcgtttacaaatcagcagggcgagcgg---g      133
              |..|| | ||||| |..|| | ||| |..||..|
B_clausii   93 -ggttgc---cagaacctt--ggttgc---cagc--gccgtactgatca      130

SEQ175     134 aaacggattttattcaatgtgtcgtttggcgccgccag--gcggaaaacg      181
              |||.....| |||||...| |..|||...|..| |||....
B_clausii  131 aaaccaggattattcccagag--gattggtttgactgttcgaggt---      175

SEQ175     182 tcgccaaactttttgaaaaaggggagcgtt--ggetggtgtcgatggccgac      229
              |||..||..| |..||| |..| ..|||..|||
B_clausii  176 tcgagaaatt-----ggacgctttcagc---aacgatttcgct-      210

SEQ175     230 tgcaaacccgcagct-----atgaaaatcaagaaggtcggcgtgtgtac      273
              |..|||..|||..| |..||| |..|..|||..|||
B_clausii  211 tacaaaaacgcgcctgccttcattattgtca--tagcttcgcgtttgcac      258

SEQ175     274 gtgacggaagt-ggtggtgatagcgtccaattttcttgagccgaaaggaa      322
              |||..| |....||..|..|..|||..|||..|...|..|
B_clausii  259 ---acggcgcgtcgactcctgcaaggctccctttttaaggaaattggcga      305

SEQ175     323 cgagcg-agc-----agcgaggggcgaca-----gcaggcggctact      358
              ||..| || |..||..|..|| |..||..|||
B_clausii  306 cgttctcagccggtttgcgccagacaacacagttaatgaagtcagcttct      355

SEQ175     359 atggggatccattcccatcgggcaagatcagaacc--accaatato--c      404
              |..||..|..| |..|..||..| |..||..|
B_clausii  356 -----cgttctcctt---gctggtttgaaaacggacggttgactgc      393

```





## **EXHIBIT 3**

# CLUSTAL W (1.83) multiple sequence alignment

Sequence format is Pearson

Sequence 1: SEQ176 164 aa

Sequence 2: G\_kaustophilus 164 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 75

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05213563.dnd]

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:2667

Alignment Score 773

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05213563.aln]

```

SEQ176      MINRVILVGRLTRDPFELRYTPSGVAVATFTLAVNRPFNTQSYENQEGRRVYVTEVVADSV 60
G_kaustophilus MINRVILVGRLTRDPFELRYTPSGVAVATFTLAVNRPFNTQ----- 40
*****

SEQ176      QFLEPKGTSEQRGATAGGYQGGERETDFIQCVVWRRQAENVANFLKKGSLAGVDGRLQTR 120
G_kaustophilus -----QGERETDFIQCVVWRRQAENVANFLKKGSLAGVDGRLQTR 80
*****

SEQ176      -----GDPFFPGQDQNHQYNEKGF 140
G_kaustophilus SYENQEGRRVYVTEVVADSVQFLEPKGTSEQRGATAGGYGDPFFPGQDQNHQYNEKGF 140
*****

SEQ176      GRIDDDPFANDGQPIDISDDDLPF 164
G_kaustophilus GRIDDDPFANDGQPIDISDDDLPF 164
*****

```

# CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ176 164 aa

Sequence 2: B\_cereus 170 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 57

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05263720.dnd]

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:2366

Alignment Score 364

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05263720.aln]

```

SEQ176      MINRVILVGRLTRDPELRYTPSGVAVATFTLAVNRPFNTQSYENQEGRR--VYVTEVVAD 58
B_cereus    MMNRVILVGRLT KDPLRYTPNGVAVATFTLAVNRAFANQQGEREADFINCVIWRKQAE 60
             *:*****:****:*****:*****:*****:*****:*****:*****:*****:

```

```

SEQ176      SVQFLEPKGTSEQRGATAGGYQQGERETDFIQCVVWRRQAENVANFLKKGSLAGVDG--- 115
B_cereus    VANYLKKGSLAGVDGRLQTRNYEQ---DGKRVYVTEVLAESVQ-FLEPRNGGGGEQRGSF 116
             .::*: . : * *: * : * . **,* **: . .* :

```

```

SEQ176      RLQTRGDPF-----PFGQDQNHQYPNEKGFGRIDDDEFFANDGQFIDISDDDLPF 164
B_cereus    NQQPSGAGFGNQGSNPFQSSNSGNQNSGFTK-NDDPFNSVVGQPIDISDDDLPF 170
             . *. * * ***** .:.* : :*****:*****

```

# CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ176 164 aa

Sequence 2: B\_anthraxis 176 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 56

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05321442.dnd]

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:2356

Alignment Score 337

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05321442.aln]

```
SEQ176      ---MINRVILVGRLTRDPELRYTPSGVAVATFTLAVNRPFTNQSYENQEGRR--VYVTEV 55
B_anthraxis MVRLMNRVILVGRLTKDPLRYTPNGVAVATFTLAVNRAFANQQGEREADFNCVIWRKQ 60
              :*****:***:*****.*****.***. **.: . * :

SEQ176      VADSVQFLEPKGTSEQRGATAGGYQGERETDFIQCVVWRRQAENVANFLKKGSLAGVDG 115
B_anthraxis AENVANYLKKGSLAGVDGRLQTRNYEGQ---DGKRVYVTEVLAEVSVQ-FLEPRNGGGEQR 116
              . : .:***: . : * *: * : * . **.* **: . .* :

SEQ176      ---RLQTRGDPF-----PFGQDQNHQYPNEKGFG--RIDDDPFANDGQPIDISDDDLPF 164
B_anthraxis GSFNQQPSGAGFGNQSSNPFGQSSNSGNQGNQNSGFTKNDDPFPSNVGQPIDISDDDLPF 176
              . *. * * *****. .:*. . :****:* *****
```

CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ176 164 aa

Sequence 2: B\_thuringiensis 173 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 56

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05344058.dnd]

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:2365

Alignment Score 351

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05344058.aln]

```
SEQ176      MINRVILVGRLTRDPFELRYTPSGVAVATFTLAVNREFTNQSVENQEGRR--VYVTEVVAD 58
B_thuringiensis MMNRVILVGRLTKDPLRYTPNGVAVATFTLAVNRAFANQQGEREADFINCVIWRKQAE 60
*:*****:***:*****:*****:*****:***:*. *.: . * : . :

SEQ176      SVQFLEPKGTSEQRGATAGGYQGERETDFIQCVVWRQAENVANFLKKGSLAGVDG--- 115
B_thuringiensis VANYLKKGSLAGVDGRLQTRNYEGQ---DGKRVYVTEVLAESVQ-FLEPRNGGGEQGRGSF 116
.:**:. . : * *: * : * . **. * **: . . * :

SEQ176      RLQTRGDFF-----PFGQDQNHQYPNEKGFG--RIDDDPFANDGQPIDISDDDLFF 164
B_thuringiensis NQOPSGAGFGNQSSNPFQSSNSGNGQGNQNSGFTKNDDPFSSNVGQPIDISDDDLFF 173
. * . * * *****. . : * . :*****: *****
```

```
Sequence format is Pearson
Sequence 1: SEQ176                      164 aa
Sequence 2: B_weihenstephanensis       173 aa
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned. Score: 56
Guide tree      file created:    [/ebi/extserv/clustalw-work/interactive/clustalw-
20060731-05393321.dnd]
Start of Multiple Alignment
There are 1 groups
Aligning...
Group 1: Sequences:   2           Score:2361
Alignment Score 349
CLUSTAL-Alignment file created    [/ebi/extserv/clustalw-work/interactive/clustalw-
20060731-05393321.aln]

SEQ176                MINRVILVGRLTRDPELRYTPSGVAVATFTTLAVNREFFTNQSYENQEGRR--VYVTEVVAD 58
B_weihenstephanensis MMNRVILVGRLTKDPDLRYTPNGVAVATFTTLAVNRAFANQQGGEREADFNCVIWRKQAE 60
*:*****:*:*****.*****.*:*. *. : . * : . :

SEQ176                SVQFLEPKGTSEQRGATAGGYQGERETDFIQCVVWRRQAENVANFLKKGSLAGVDG--- 115
B_weihenstephanensis VANYLKKGSLAGVDGRLQTRNYDGQ---DGKRIVYVTEVLAESVQ-FLEPRNGGGEQRGSF 116
.:*: . : * *: * : * . **.* **: . .* :

SEQ176                RLQTRGDPF-----PFGQDQNHQYPNEKGFG--RIDDDPFANDGQPIDISDDDLPF 164
B_weihenstephanensis NQFPSCAGFGNQSSNPFGQSSNSGNQGNGSNGFTKNDDPFPSNVGQPIDISDDDLPF 173
. *. * * ****.* .:*. :*****
```

```

CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson
Sequence 1: SEQ176          164 aa
Sequence 2: B_subtilis     172 aa
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned. Score: 59
Guide tree      file created:  [/ebi/extserv/clustalw-work/interactive/clustalw-
20060731-05422505.dnd]
Start of Multiple Alignment
There are 1 groups
Aligning...
Group 1: Sequences:  2      Score:2402
Alignment Score 388
CLUSTAL-Alignment file created  [/ebi/extserv/clustalw-work/interactive/clustalw-
20060731-05422505.aln]

```

```

SEQ176      MINRVILVGRLTRDPRLRYTPSGVAVATFTLAVNRPFNTQSYENQEGRRVYVT--EVVAD 58
B_subtilis  MLNRVVLVGRLTQDPRLRYTPNGAAVATFTLAVNRFTNTQSGEREADFINCVTWRRAEN 60
*:***:*****:*****:*,*****:*****:***** *: . . . . . :

SEQ176      SVQFLEPKGTSEQRGATAGGYQGERETDFIQCVVWRRQAENVANFLKK-----GSL 110
B_subtilis  VANFLKKGSLAGVDGRLQTRNYENQQGQ---RVFVTEVQAESVQFLEPKNGGGSGSGGYN 117
.:**:. . : * .: .: . : . * . . . . :

SEQ176      AGVDGRLQTRG---DPFPFGQDQNHQYPNEKGFGRIDDDPFANDGQPIDISDDDLPF 164
B_subtilis  EGNSSGGQYFGGGQNDNPFGGNQNNQRRNQG--NSFNDDPFANDGKPIDISDDDLPF 172
* . * * * : *** :*** * *: . :*****:*****

```



# CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ176 164 aa

Sequence 2: B\_licheniformis 170 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 60

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05435846.dnd]

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:2430

Alignment Score 399

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05435846.aln]

```

SEQ176      MINRVILVGRLTRDPELRYTPSGVAVATFTLAVNRPFTNQSYENQEGRR--VYVTEVVAD 58
B_licheniformis MLNRVVLVGRLT KDPELRYTPSGAAVATFTLAVNRTFTNQGGEREADFINCVVWRRQAEN 60
*:***:*****:*****:*****:*****:*****:*.***. *: . * . . :

SEQ176      SVQFLEPKGTSEQRGATAGGYQQGERETDFIQCVVWRRQAENVANFLKKGSLAGVDGRLQ 118
B_licheniformis VANFLKKGSLAGVDGRLQTRSYENQQGQ--RVYVTEVQAESVQFLEPKGGSGSGGYSG 117
.:**:. . : * .: .: : * . ***.* : **. :* .*

SEQ176      TRG-----DFFPFGQDQNHQYFNEKGFGGRIDDDPFANDGQPIDISDDDLPF 164
B_licheniformis GQGGQHFGGGQNEPAPFGGSQNNQNRNQ--NSFNDDPFANDGKPIDISDDDLPF 170
:* .: .: * *** .**:* *: . :*****:*****

```

# CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ176 164 aa

Sequence 2: B\_halodurans 168 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 54

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05480066.dnd]

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:2356

Alignment Score 350

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05480066.aln]

```

SEQ176      MINRVILVGRLTRDPELRYTPSGVAVATFTLAVNRPFTNQSYENQEGRR-VYVTEVVADS 59
B_halodurans MLNRVVLVGRLTRDPELRYTPNGVAVANFTLAVNRPFSNQGGEREADFINCVWRKQAE 60
              *:***:*****:*****:*****:*****:***. *: . . . * . *:..

SEQ176      VQFLEPKGTSEQRGATAGGYQGERETDFIQCVVWRQAENVANFLKKGSLAGVDGRLQT 119
B_halodurans VANYLKKGSG--LAGVDGRIQTRSYDNNEGRRVFVTEVMAESVQFLEPRGSQSQGGSNVDN 118
              *      **:      *. . . . . :. :. : :.* . **.* : :** : ...:.

SEQ176      RG--DF-FPFGQDQNHQYPNEKG--FGRIDDDPFANDGQPIDISDDDLPF 164
B_halodurans FGGGSPNNPMGGNDFGQQSGGSGRQSGGFSEDPFANDGKPIDISDDDLPF 168
              * . * *:* : * . . * * :.:*****:*****

```

CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ176 164 aa

Sequence 2: B\_clausii 161 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 54

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05513189.dnd]

Start of Multiple Alignment

There are 1 groups

Aligning...

Group 1: Sequences: 2 Score:2277

Alignment Score 542

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060731-05513189.aln]

```
SEQ176      MINRVILVGRLTRDPELRYTPSGVAVATFTLAVNRPFTNQSYENQEGRRVYVTEVVADSV 60
B_clausii   MLNRVVLVGRLTRDPELRETPNGVAVANFTLAVNRPFSNQ----- 40
*:***:*****:*.*****:*****:*
```

```
SEQ176      QFLEPKGTSEQRGATAGGYQGERETDFIQCVVWRQAENVANFLKKGSLAGVDGRLQTR 120
B_clausii   -----QGEREADFINCVVWRKPAENVANFLKKGSLAGVDGRVQTR 80
          *****:***:*****: *****:***
```

```
SEQ176      -----GDPFPFGQDQNHQYP---NEKGFGR----- 142
B_clausii   SYDNNEGRRVFVTEIVAESVQFLEPRNSQSSGNNPGFDQYGAGNQSGNQGTGGSRSNG 140
          .... * : : * *      *: **:
```

```
SEQ176      IDDDPFFANDGQPIDISDDDLPF 164
B_clausii   YDNDPFSNDG-SIDISDDDLPF 161
*:***:*** .*****
```